

MEASUREMENT THEORY AND THE FOUNDATIONS OF UTILITARIANISM

by

John A. Weymark



Working Paper No. 05-W07

April 2005

DEPARTMENT OF ECONOMICS
VANDERBILT UNIVERSITY
NASHVILLE, TN 37235

www.vanderbilt.edu/econ

Measurement Theory and the Foundations of Utilitarianism

JOHN A. WEYMARK

*Department of Economics, Vanderbilt University,
VU Station B #35189, 2301 Vanderbilt Place,
Nashville, TN 37235-1819, U.S.A.*
(e-mail: john.weymark@vanderbilt.edu)

April 2005

Abstract. This article reconsiders the Harsanyi–Sen debate concerning whether Harsanyi is justified in interpreting his Aggregation and Impartial Observer Theorems as providing axiomatizations of utilitarianism. Sen’s criticism and its formalization by Weymark are based on the claim that von Neumann–Morgenstern utility theory is ordinal, whereas Harsanyi’s utilitarian conclusions require cardinal utility. Proposals for overcoming Sen’s objection that appeal to formal measurement theory are considered. It is argued that one of these proposals due to Broome and Risse rightly points to a feature of expected utility theory that was ignored by Sen and Weymark, but that this proposal does not provide a normatively compelling justification for cardinal utility. The other proposal due to Broome is shown to make use of a strength of preference relation in addition to the axioms of expected utility theory.

Keywords and Phrases: expected utility, measurement theory, utilitarianism, Harsanyi, von Neumann, Morgenstern.

JEL Classification Numbers: D63, D71, D81.

1. Introduction

John Harsanyi's 1953 and 1955 articles in the *Journal of Political Economy* [Harsanyi (1953, 1955)] are among the most important and influential contributions to social choice theory and welfare economics. In these articles, Harsanyi offered a rational choice foundation for utilitarianism. Specifically, Harsanyi used the recently developed expected utility theory of von Neumann and Morgenstern (1944) to provide two axiomatizations of utilitarianism. In Weymark (1991), I referred to these results as Harsanyi's Aggregation and Impartial Observer Theorems.

In Harsanyi's Aggregation Theorem, individual and social preferences on the set of lotteries generated by a finite set of sure outcomes are assumed to satisfy the axioms of expected utility theory. Furthermore, two lotteries are socially indifferent if every individual is indifferent between them—the familiar Pareto Indifference condition. With these assumptions, Harsanyi has shown that if the preferences are represented by von Neumann–Morgenstern utility functions, then the social utility function is an affine function of the individual utility functions. Hence, lotteries are socially ranked according to a “weighted utilitarian” rule.

In Harsanyi's Impartial Observer Theorem, a hypothetical impartial observer determines a social ranking of the alternatives by imagining himself as having an equal chance of being any individual in society. The individual preferences over lotteries are assumed to satisfy the expected utility axioms, as are the observer's preferences over the extended lotteries in which both his identity and the outcome of the actual lottery are uncertain. Provided that the observer is sympathetic to the interests of the individuals and that his preferences over extended lotteries are represented by a von Neumann–Morgenstern utility function, lotteries are socially ranked according to their average utility; i.e., by an “average utilitarian” rule.

Sen (1976) has argued that neither of Harsanyi's theorems are axiomatizations of utilitarian rules because he has misapplied expected utility theory. In particular, Sen has argued that, contrary to what many believe, von Neumann–Morgenstern expected utility theory is an ordinal theory and, therefore, any increasing transform of a von Neumann–Morgenstern utility function is a satisfactory representation of an individual's preference relation. However, utilitarianism requires a cardinal theory of utility and so, according to Sen, Harsanyi was not justified in giving his theorems utilitarian interpretations. Sen's informal discussion of these issues has been formalized by Wey-

mark (1991). Broome (1997) calls this argument the “standard objection” to Harsanyi’s theorems. A particularly clear presentation of this objection from someone who endorses its conclusions may be found in Roemer (1996b, Chapter 4).

In the Sen–Weymark critique of Harsanyi, von Neumann–Morgenstern utility theory is identified with the utility representation of a binary preference relation over lotteries satisfying the expected utility axioms. However, as Risse (2002) has recently reminded us, von Neumann and Morgenstern (1944, Chapter 1, Section 3) regard their main contribution to utility theory to be the “discovery” of a “natural” operation (the convex combination operator used to form a probability distribution over outcomes) that allows them to put the measurement of utility on as firm a foundation as the measurement of, say, heat. It is this operation in conjunction with the preference relation that makes utility cardinally measurable.

Von Neumann and Morgenstern’s axiomatization of “measurable” utility is now regarded as being one of the pioneering contributions to what has come to be known as the representational theory of measurement.¹ Broome (1991, 1997) and Risse (2002) have in various ways made use of measurement theory in order to evaluate the extent to which Harsanyi’s theorems can be used to provide support for some form of utilitarianism. Broome (1997) and Risse (2002) have argued that the Sen–Weymark objection to Harsanyi can be overcome by taking proper account of von Neumann and Morgenstern’s convex combination operator. Broome (1991), and latter Risse (2002), invoke arguments concerning the weighing of goods across states of nature to help justify the use of von Neumann–Morgenstern representations in Harsanyi’s theorems.

In this article, I argue that even if full account is taken of the role that convex combination operators play in von Neumann and Morgenstern’s theory, the sense in which von Neumann–Morgenstern utility is cardinal is not a sense that has normative content. Rather, the selection of a von Neumann–Morgenstern representation of an individual’s preference (instead of some non-affine transform of this function) is based on pragmatic considerations that may be compelling when describing choice behaviour in the presence of uncertainty, but that are not compelling when the objective is to mea-

¹See, for example, Krantz, Luce, Suppes, and Tversky (1971) or Roberts (1979) for introductions to measurement theory. Ellingsen (1994) provides an insightful account of the history of cardinal utility from a measurement-theoretic perspective.

sure an individual’s well-being, as is required by normative theories such as utilitarianism. I also argue that the way in which Broome (1991) circumvents this problem is by implicitly supplementing the preference relation of von Neumann and Morgenstern with a second relation that quantitatively measures utility differences. I therefore conclude that the “standard objection” to Harsanyi, while employing an incomplete account of von Neumann–Morgenstern expected utility theory, can be extended to take the complete description of their utility theory into account. The cardinality of utility necessary for Harsanyi’s utilitarian interpretation of his theorems to be justifiable must be found elsewhere. My arguments apply to both of Harsanyi’s theorems, so, for concreteness, I focus my discussion on his Aggregation Theorem.

In his theorems, Harsanyi uses the axiomatization of expected utility theory due to Marschak (1950), rather than the original axiomatization of von Neumann and Morgenstern (1944). Marschak’s version of expected utility theory is described in Section 2. In Section 3, I discuss how Harsanyi’s 1953 and 1955 articles provided a connection between expected utility theory and welfare economics. I also formally state his Aggregation Theorem. In Section 4, I present a summary of the Sen–Weymark critique of Harsanyi. An introduction to the representational theory of measurement is provided in Section 5. In Section 6, I describe how expected utility theory can be formalized using the kinds of structures employed in measurement theory and argue that this formalization does not provide a normatively compelling justification for the cardinality of utility. I then consider Broome’s and Risse’s arguments in Section 7. I close with some concluding remarks in Section 8.

2. Expected Utility Theory

Harsanyi (1955) models uncertainty using lotteries over a finite set of sure outcomes. Formally, the set of *sure outcomes* is $X = \{x_1, \dots, x_M\}$, where $M \geq 2$. A (simple) *lottery* $p = (p_1, \dots, p_M)$ specifies, for each outcome $x_m \in X$, the probability p_m of obtaining this outcome. The set of all lotteries \mathcal{L} on X is the $(M - 1)$ -dimensional unit simplex; i.e., the set of all $p \in \mathbb{R}_+^M$ for which $\sum_{m=1}^M p_m = 1$. Letting $e^m = (e_1^m, \dots, e_M^m)$, where $e_i^m = 1$ if $i = m$ and $e_i^m = 0$ otherwise, the sure outcome x_m can be equivalently described by the lottery e^m that assigns probability one to the outcome x_m .

A *weak preference relation* is a binary relation \succeq on \mathcal{L} . The corresponding

strict preference \succ and indifference \sim relations are defined by setting, for all $p, q \in \mathcal{L}$, $p \succ q \leftrightarrow [p \succeq q \text{ and } \neg(q \succeq p)]$ and $p \sim q \leftrightarrow [p \succeq q \text{ and } q \succeq p]$. The binary relation \succeq is *complete* if for all $p, q \in \mathcal{L}$, $p \succeq q$ or $q \succeq p$ and it is *transitive* if for all $p, q, r \in \mathcal{L}$, $[p \succeq q \text{ and } q \succeq r] \rightarrow p \succeq r$.

For all $p, q \in \mathcal{L}$ and all $\alpha \in [0, 1]$, the convex combination $[\alpha p + (1 - \alpha)q]$ defines a new lottery. This lottery can be given two interpretations. First, because $[\alpha p + (1 - \alpha)q]$ is simply a vector in the $(M - 1)$ -dimensional unit simplex, it can be interpreted as being a simple lottery. Second, it can be interpreted as being a two-stage compound lottery. A *compound lottery* is a lottery that has lotteries as prizes. In the case of $[\alpha p + (1 - \alpha)q]$, in the first stage, with probability α the outcome is the simple lottery p and with probability $1 - \alpha$ the outcome is the simple lottery q .²

There are many essentially equivalent ways of axiomatizing expected utility theory when the set of alternatives are lotteries. The first, of course, is the axiomatization of von Neumann and Morgenstern (1944). Harsanyi (1955) uses Axioms I, II, III', and IV of Marschak (1950), rather than the axioms originally proposed by von Neumann and Morgenstern. These four Marschak axioms are: Ordering, Continuity, Nondegeneracy, and Independence.

Ordering. \succeq is complete and transitive.

Continuity. For all $p, q, r \in \mathcal{L}$ for which $p \succ q \succ r$, there exists an $\alpha \in (0, 1)$ such that $[\alpha p + (1 - \alpha)r] \sim q$.

Nondegeneracy. There exist four distinct lotteries in \mathcal{L} that are mutually nonindifferent.

Independence. (i) For all $p, q \in \mathcal{L}$ and all $\alpha \in (0, 1)$, if $p \sim q$, then $p \sim [\alpha p + (1 - \alpha)q]$. (ii) For all distinct $p, q, r \in \mathcal{L}$ and all $\alpha \in (0, 1)$, if $p \sim q$, then $[\alpha p + (1 - \alpha)r] \sim [\alpha q + (1 - \alpha)r]$.

²Strictly speaking, a compound lottery is not in \mathcal{L} , so when the compound lottery interpretation of $[\alpha p + (1 - \alpha)q]$ is adopted, preferences need to be defined on the set of lotteries, both simple and compound, that can be generated by the set of outcomes X . Versions of expected utility theory that employ compound lotteries assume that any compound lottery is indifferent to the simple lottery that assigns the same probability to each of the outcomes in X as does the compound lottery. With this assumption, there is no loss of generality in simply defining preferences on \mathcal{L} , as is done here. Note that this "reduction of compound lotteries" assumption rules out the possibility that an individual has an intrinsic taste for gambling. See, for example, Harsanyi (1987).

Ordering says (i) that the decision-maker is able to compare any two lotteries and either say that they are indifferent to each other or that one is strictly preferred to the second and (ii) that the preference \succeq does not exhibit a cycle over any triple of lotteries unless these alternatives are all indifferent to each other. Continuity says that if three lotteries can be strictly ranked, then the middle-ranked lottery is indifferent to some probability mixture of the other two lotteries. Nondegeneracy requires that there be at least four distinct indifference classes of lotteries. In combination with the other axioms, Nondegeneracy implies that there are in fact a infinite number of indifference classes. Independence says that if two lotteries p and q are indifferent to each other, then (i) so are all the lotteries on the line segment joining them and (ii) any probability mixture between p and some third lottery r is indifferent to the same probability mixture between q and r .³

The axioms of expected utility theory are usually interpreted as being necessary conditions for a preference over risky alternatives to qualify as being rational. In behavioural interpretations of the theory, preferences are defined from choice. For example, the statement that $p \succ q$ is taken to mean that only p is chosen when confronted with a choice from the set $\{p, q\}$. With this interpretation, the choices made from pairs of lotteries are not necessarily meant to be descriptive of a decision-maker's actual behaviour when confronted with these options; it could instead correspond to the behaviour of an ideally rational decision-maker. Regardless of whether preferences are a primitive of the model or are defined from choice, the axioms of expected utility theory need to be provided with some justification as to why rational preferences must satisfy them. Detailed discussions of some of the justifications that have been offered in support of expected utility axioms may be found in Luce and Raiffa (1957, Chapter 2) and Broome (1991, Chapter 5).

However, in order for expected utility theory to provide a normative foundation for utilitarianism (or, at least, for weighted utilitarianism), a different justification for the axioms of expected utility theory is needed. Broome (1991) and Risse (2002) have argued that the requisite justification has been provided if it can be shown that the concept of well-being (or goodness) that constitutes the "utility" that enters into utilitarian calculations entails a *betterness binary relation* on the set of alternatives that conforms to the

³By not requiring r to be distinct from p , part (i) of this axiom can be subsumed in part (ii).

axioms of expected utility theory. This betterness relation is defined by saying that alternative a is weakly better than alternative b if and only if the well-being associated with a is at least as great as that associated with b . Risse models uncertainty using lotteries, as is done here, and has an idealized desire satisfaction account of well-being. Broome, on the other hand, uses the state-contingent alternatives approach to uncertainty and identifies utility with goodness.⁴ Both Broome and Risse, in the context of their respective theories, argue that the betterness relation described above should conform to the axioms of expected utility theory, or at least that these axioms are reasonable constraints on this relation.

A function $U: \mathcal{L} \rightarrow \mathbb{R}$ is a *utility function* representing \succeq if

$$U(p) \geq U(q) \leftrightarrow p \succeq q, \quad \forall p, q \in \mathcal{L}. \quad (1)$$

A function $f: \mathbb{R} \rightarrow \mathbb{R}$ is an *increasing transform* if $f(s) \geq f(t)$ for all $s, t \in \mathbb{R}$. If U is a utility function representing \succeq , then so is any increasing transform of U . A utility function that is unique up to an increasing transform is said to be *ordinal*.

A function $U: \mathcal{L} \rightarrow \mathbb{R}$ is a *von Neumann–Morgenstern utility function* representing \succeq if (1) is satisfied and if

$$U(p) = \sum_{m=1}^M p_m U(e^m), \quad \forall p \in \mathcal{L}. \quad (2)$$

A von Neumann–Morgenstern utility function is a linear function on the set of lotteries \mathcal{L} . Because the lottery e^m corresponds to receiving the outcome x_m for certain, the utility assigned to any lottery p by a von Neumann–Morgenstern utility function is the expected value of the utilities assigned to the sure outcomes in the set X .

A function $f: \mathbb{R} \rightarrow \mathbb{R}$ is an *increasing affine transform* if there exist real numbers a and b with $b > 0$ such that $f(t) = a + bt$ for all $t \in \mathbb{R}$. If the image $U(\mathcal{L})$ of the von Neumann–Morgenstern utility function U is a nondegenerate interval, then $V: \mathcal{L} \rightarrow \mathbb{R}$ is also a von Neumann–Morgenstern utility function if and only if V is an increasing affine transform of U .⁵ In other words, the

⁴In the state-contingent alternative approach to uncertainty, alternatives are acts that map states of nature into consequences. The best-known formalization of this approach is due to Savage (1954).

⁵The image of U is the set $U(\mathcal{L}) = \{t \in \mathbb{R} | U(p) = t \text{ for some } p \in \mathcal{L}\}$. If all lotteries are indifferent to each other, then any utility function representing \succeq is trivially a von Neumann–Morgenstern utility function.

utility function U is unique up to an increasing affine transform. A utility function that is unique up to an increasing affine transform or any smaller set of transforms is said to be *cardinal*.

Theorem 1 is the version of the *Expected Utility Theorem* used by Harsanyi (1955).⁶

Theorem 1. *If a binary relation \succeq on \mathcal{L} satisfies Ordering, Continuity, and Independence, then it can be represented by a von Neumann–Morgenstern utility function U . If, in addition, \succeq satisfies Nondegeneracy, then the functional form of a von Neumann–Morgenstern utility function representing \succeq is unique up to an increasing affine transform.*⁷

Theorem 1 does not say that in order for a utility function to represent a preference relation \succeq that satisfies Ordering, Continuity, and Independence, it must be a von Neumann–Morgenstern utility function. A utility representation only needs to satisfy (1), not both (1) and (2). As a consequence, any increasing transform of a von Neumann–Morgenstern utility function also represents \succeq .

This point has been poorly understood in spite of the fact that many of the early discussions of von Neumann and Morgenstern’s theory, such as those of Arrow (1951) and Baumol (1951, 1958), explicitly noted this fact. For example, Arrow (1951, p. 10) says that “... instead of using the utility scale found by von Neumann and Morgenstern, we could use the square of that scale; then behavior is described by saying that the individual seeks to maximize the expected value of the square root of his utility.”⁸

What justifications have been offered for singling out von Neumann–Morgenstern utility functions from the class of all representations of \succeq ? The main justification has been that of convenience. It is simply more convenient

⁶Harsanyi (1955) does not explicitly state this theorem. He instead refers the reader to Marschak (1950). In my statement of the Expected Utility Theorem, I have clarified the fact that Nondegeneracy is not needed to establish the existence of a von Neumann–Morgenstern utility function.

⁷The only role of the nondegeneracy axiom in Theorem 1 is to rule out the case in which all lotteries are indifferent to each other. Thus, this axiom can be weakened so that it only requires the existence of two nonindifferent lotteries. In the version of the expected utility theorem presented in Weymark (1991), I neglected to explicitly rule out the possibility of universal indifference when considering the uniqueness properties of von Neumann–Morgenstern utility functions. I also used stronger forms of the continuity and independence axioms.

⁸Arrow is implicitly assuming that utility is always nonnegative in his example.

to describe an individual's preferences using an expected utility representation than by some non-affine transform of such a function. For example, calculations are simpler when a von Neumann–Morgenstern utility function is used rather than a utility function of the form described by Arrow. This rationale for the use of von Neumann–Morgenstern representations has been offered by Marschak (1950, pp. 131–132), Arrow (1951, p. 10), Friedman and Savage (1952, pp. 471–472), and Alchian (1953, p. 39), among others.

Another justification for the use of a von Neumann–Morgenstern representation is that it provides a parsimonious description of preferences. By simply knowing the values assigned by a von Neumann–Morgenstern utility function to the sure outcomes, using (2) it is possible to infer how all of the lotteries in \mathcal{L} are ranked. This advantage of von Neumann–Morgenstern representations has been noted by, for example, Friedman and Savage (1948, p. 292), Ellsberg (1954, p. 537), and Baumol (1958, p. 668). However, this parsimony property is not unique to von Neumann–Morgenstern utility functions, as Arrow's example demonstrates.

3. Harsanyi's Linking of Expected Utility Theory and Welfare Economics

In the decade following the publication of the first edition of von Neumann and Morgenstern's monograph, many commentators considered whether von Neumann–Morgenstern utility functions had any relevance for welfare economics. The discussion of this issue was often linked to the related question of how, if at all, von Neumann and Morgenstern's "cardinal" utility functions differed from the cardinal utility functions of neoclassical economists such as Jevons. A distinguishing feature of a neoclassical utility function is that the marginal utility of income decreases with an increase in income.⁹

There was widespread, but not universal, agreement that von Neumann and Morgenstern's expected utility theory has little or no welfare significance.

⁹Mandler (1999) provides an illuminating discussion of the history of utility theory that focuses on the ordinalist-cardinalist controversy. As Ellingsen (1994) has documented, a great deal of confusion has resulted from the failure to distinguish between conceptually distinct methods for constructing a utility index. For example, while both Edgeworth (1881) and Pareto (1906) have cardinal theories of utility, Edgeworth regards a unit of utility as being a just perceptible increment of pleasure, whereas Pareto's theory is based on directly comparing differences in pleasure (*ofelimità*), as in the statement that the increase in pleasure from substituting a for b exceeds that from substituting c for d .

For example, Friedman and Savage (1952, p. 473) said that simply because individuals “. . . act as if they are maximizing the expected value of a function unique except for origin and unit of measurement has, in and of itself, no welfare implications at all . . .”. Similarly, Arrow (1951, p. 10) said that the expected utility theorem “. . . has nothing to do with welfare considerations, particularly if we are interested primarily in making a social choice among alternative policies in which no random elements enter.”

The same view was expressed in widely-read and influential articles by Alchian (1953) and Ellsberg (1954) that attempted to convey the essentials of von Neumann–Morgenstern expected utility theory and its implications for the measurability of utility to the typical economist of the time who did not have the mathematical skills to follow the rather technical arguments in von Neumann and Morgenstern (1944) and Marschak (1950). Ellsberg (1954, p. 556) asserted that a von Neumann–Morgenstern utility index (function) does not “. . . seem to be of any relevance to welfare evaluations (whereas a Jevonsian index might be).” Alchian (1953, p. 50) is even more emphatic, saying that von Neumann–Morgenstern expected utility “. . . has literally nothing to do with individual, social or group welfare, whatever the latter is supposed to mean.”

Moreover, Friedman and Savage (1948) had argued that (in the words of Harsanyi (1953, p. 434)) “[i]n the theory of risk-taking . . . , increasing marginal utility is to be assumed to prevail over a considerable range, in view of people’s willingness in the case of gambling to pay a price far above the actuarial value for a small chance of a large gain.” Thus, von Neumann–Morgenstern utility should not be equated with the utility of the neoclassical school.

Harsanyi’s objective in Harsanyi (1953) was to dispute the claim that von Neumann–Morgenstern expected utility theory had no relevance for welfare economics while at the same time allowing for the possibility that Friedman and Savage were correct about how marginal utility varies with income. For Harsanyi, welfare judgments are the impersonal preferences expressed by an impartial observer who orders social alternatives based on a sympathetic but impartial concern for the interests of everyone in society. Specifically, the impartial observer engages in a thought experiment in which he imagines having an equal chance of being anyone in society, complete with that person’s preferences and objective circumstances. It then follows, or so Harsanyi (1955, p. 316) argued, that “. . . without any additional ethical postulates that an individual’s impersonal preferences, if they are rational, must satisfy

Marschak's axioms and consequently must define a cardinal social welfare function equal to the arithmetic mean of the utilities of all individuals in the society ...".¹⁰ In Weymark (1991), I called this result Harsanyi's Impartial Observer Theorem.¹¹

The original discussions of the Impartial Observer Theorem in Harsanyi (1953, 1955) are quite informal. The first formal presentation of this result appears in Harsanyi (1977b, Chapter 4), although even here, Harsanyi glosses over some of the details. In Weymark (1991), I provided what I believe is a complete formal statement of Harsanyi's Impartial Observer Theorem and its proof.

In Harsanyi (1955), a way of linking expected utility theory with welfare economics was introduced that is conceptually distinct from his Impartial Observer Theorem, what in Weymark (1991) I have called Harsanyi's Aggregation Theorem. In this theorem, Harsanyi assumes that every individual in society has preferences on \mathcal{L} that satisfy Marschak's version of the expected utility axioms. There is also a social preference on \mathcal{L} that satisfies these axioms. In Harsanyi's favoured interpretation, a social (or moral) preference is the preference of one of the members of society when he evaluates the alternatives in \mathcal{L} from an impersonal perspective. Social and individual preferences are related by the requirement that if everyone is indifferent between two lotteries, then they are also socially indifferent. This is simply the familiar *Pareto Indifference* axiom of welfare economics. Harsanyi's Aggregation Theorem shows that if these assumptions are satisfied and if the individual and social preferences are represented by von Neumann–Morgenstern utility functions, then the social utility function is an affine combination of the individual utility functions.

Harsanyi has interpreted both the Impartial Observer Theorem and the Aggregation Theorem as being theorems about (weighted) utilitarianism.

¹⁰I have suppressed a footnote in which Harsanyi notes that computing the average of utilities yields the same ranking of social alternatives as taking the sum (when the population is fixed). By "a social welfare function," Harsanyi means a utility function representing the impartial observer's preferences.

¹¹Although implicit, the Impartial Observer Theorem is not stated explicitly in Harsanyi (1953). In that article, the social alternatives are distributions of income, although in the later formulations of the Impartial Observer Theorem in Harsanyi (1955, 1977b), alternatives are given broader interpretations. A version of the Impartial Observer Theorem for income distributions had previously appeared in Vickrey (1945), but, according to a 1996 interview with Harsanyi, he was unaware of this fact until he went to Stanford in 1956. See d'Aspremont and Hammond (2001, pp. 392–393).

Some of the criticisms that have been levelled against these interpretations will be taken up in the next section. For concreteness, and because it is simpler to state formally, I will focus on the Aggregation Theorem, although the main points I am making apply to both results. In order to evaluate the validity of Harsanyi's interpretation of the Aggregation Theorem, it is necessary to state it formally.

There are n individuals, indexed by $i = 1, \dots, n$. Individual i has a preference relation \succeq_i on \mathcal{L} . There is also a social preference relation \succeq on \mathcal{L} . The only link between the individual and social preferences is provided by the Pareto Indifference condition.

Pareto Indifference. For all $p, q \in \mathcal{L}$, if $p \sim_i q$ for all $i = 1, \dots, n$, then $p \sim q$.

Theorem 2 is a formal statement of Harsanyi's Aggregation Theorem without his redundant Nondegeneracy axiom.

Theorem 2. *Suppose that \succeq_i , $i = 1, \dots, n$, and \succeq are binary relations on \mathcal{L} that satisfy Ordering, Continuity, and Independence and also suppose that Pareto Indifference is satisfied by these relations. Let U_i be a von Neumann–Morgenstern utility representation of \succeq_i , $i = 1, \dots, n$, and U be a von Neumann–Morgenstern utility representation of \succeq . Then, there exist real numbers a_i , $i = 1, \dots, n$, and b such that*

$$U(p) = \sum_{i=1}^n a_i U_i(p) + b, \quad \forall p \in \mathcal{L}. \quad (3)$$

An implication of (3) is that

$$U(p) \geq U(q) \leftrightarrow \sum_{i=1}^n a_i U_i(p) \geq \sum_{i=1}^n a_i U_i(q), \quad \forall p, q \in \mathcal{L}. \quad (4)$$

Thus, p is socially weakly preferred to q if and only if a weighted sum of the individual von Neumann–Morgenstern utilities associated with p is at least as great as the corresponding sum for q . The conclusion Harsanyi draws from this observation is that alternatives are socially ranked using a weighted utilitarian rule.

The assumptions of Theorem 2 do not guarantee that there is only one vector (a_1, \dots, a_n, b) for which (3) is satisfied. Uniqueness of these coefficients obtains if it is further assumed that for each individual i , there is a pair of lotteries for which i is not indifferent, but every other individual is.¹² In Weymark (1991), this assumption is called *Independent Prospects*.

Independent Prospects. For all $i = 1, \dots, n$, there exist lotteries $p_i, q_i \in \mathcal{L}$ such that $\neg[p_i \sim_i q_i]$ and $p_i \sim_j q_i$ for all $j \neq i$.

Note that if Independent Prospects is satisfied and \succeq_i can be represented by a von Neumann–Morgenstern utility function, then \succeq_i must satisfy Nondegeneracy.

Furthermore, the assumptions of Theorem 2 do not imply that the weights a_i can all be chosen to be positive or even nonnegative. These sign restrictions can be satisfied if stronger forms of the Pareto condition are invoked. For an extended discussion of these variant forms of Harsanyi’s Aggregation Theorem, see Weymark (1991).

4. The Harsanyi–Sen Debate

The debate carried out in Harsanyi (1975, 1977a) and Sen (1970, 1976, 1977, 1986) about the ethical significance of Harsanyi’s theorems and their relationship to utilitarianism raised many issues. Here, I shall consider Sen’s argument that Harsanyi is not justified in giving his theorems a utilitarian interpretation. The material in this section draws upon Weymark (1991), to which the reader is referred for more detail.¹³

As in the previous section, the individual and social preference relations $\succeq_i, i = 1, \dots, n$, and \succeq on \mathcal{L} are fixed. A *profile of utility functions* is an n -tuple $\mathbf{U} = (U_1, \dots, U_n)$ of individual utility functions $U_i, i = 1, \dots, n$, where each U_i is defined on \mathcal{L} . Let \mathcal{U} denote the set of all profiles of utility functions for which for all $\mathbf{U} \in \mathcal{U}$ and all $i = 1, \dots, n$, U_i is a representation of \succeq_i . Further restrictions may be placed on the set of admissible utility

¹²Harsanyi (1955) implicitly made this assumption in his proof of the Aggregation Theorem. The first proof of this theorem without this implicit assumption was by Domotor (1979).

¹³See also the discussion of Harsanyi’s theorems in Mongin and d’Aspremont (1998) and Roemer (1996b).

representations. For example, only von Neumann–Morgenstern utility representations may be admissible. Let $\mathcal{U}^a \subseteq \mathcal{U}$ denote the set of *admissible* profiles of utility functions.

Weighted utilitarianism for profiles in \mathcal{U}^a requires lottery p to be socially ranked at least as good as lottery q if and only if the weighted sum of the utilities obtained from p is at least as large as the weighted sum of the utilities obtained from q , where the weights are profile independent. In Harsanyi’s Aggregation Theorem, there is only one social preference relation \succeq . Thus, \succeq is *weighted utilitarian* if there exists a weight vector $a = (a_1, \dots, a_n) \in \mathbb{R}^n$ for which for all $\mathbf{U} \in \mathcal{U}^a$,

$$p \succeq q \leftrightarrow \sum_{i=1}^n a_i U_i(p) \geq \sum_{i=1}^n a_i U_i(q), \quad \forall p, q \in \mathcal{L}. \quad (5)$$

Classical utilitarianism corresponds to the case in which all the weights are equal.¹⁴

In the context of the problem being considered by Harsanyi (fixed individual and social preference relations), Sen, in the references cited above, has argued that utility is only used to *represent* preferences in Harsanyi’s Impartial Observer and Aggregation theorems, and this is an inadequate basis on which to construct an axiomatic foundation for (weighted) utilitarianism. As noted above, I restrict attention to Harsanyi’s Aggregation Theorem, as the basic point that Sen has raised applies to both of Harsanyi’s theorems.¹⁵

Consider a profile \mathbf{U} of von Neumann–Morgenstern utility functions and suppose that this profile is in \mathcal{U}^a . By Theorem 2, we know that there exists a weight vector a such that (5) holds for the profile \mathbf{U} . In order to focus on the essentials of the argument, let me further suppose that Independent Prospects is satisfied, so that a is the only weight vector for which (5) is true for this profile.

Even if we restrict attention to von Neumann–Morgenstern utility representations of the individual preferences, Theorem 2 does not permit us to interpret \succeq as being a weighted utilitarian social preference relation. To see

¹⁴Harsanyi’s objective was to provide a choice-theoretic foundation for classical utilitarianism. However, as Mongin and d’Aspremont (1998) have argued, the kind of symmetry condition needed to justify the use of the same weight for all individuals in (5) requires a multi-profile framework in which different profiles of individual preferences and their corresponding social preferences are considered.

¹⁵See Weymark (1991) and Roemer (1996a,b) for discussions of Sen’s critique of Harsanyi’s Impartial Observer Theorem.

why, suppose that \mathcal{U}^a contains all profiles of von Neumann–Morgenstern representations of $(\succeq_1, \dots, \succeq_n)$. Consider the profile of von Neumann–Morgenstern utility functions \mathbf{U}' for which $U'_1 = 2U_1$ and $U'_i = U_i$ for all $i \neq 1$. Using \mathbf{U}' instead of \mathbf{U} in Theorem 2, it follows that there exist weights $\bar{a} = (\bar{a}_1, \dots, \bar{a}_n)$ such that

$$p \succeq q \leftrightarrow \bar{a}_1 2U_1(p) + \sum_{i=2}^n \bar{a}_i U_i(p) \geq \bar{a}_1 2U_1(q) + \sum_{i=2}^n \bar{a}_i U_i(q), \quad \forall p, q \in \mathcal{L}. \quad (6)$$

Because the image of U_1 is a nondegenerate interval, in order for (5) and (6) to describe the same social preference, we must have $\bar{a}_1 = a_1/2$ and $\bar{a}_i = a_i$ for all $i \neq 1$. Hence, the weights that need to be used to represent the social preference relation \succeq as a weighted sum of individual utilities is profile dependent. However, weighted utilitarianism requires the same weights (up to a common factor of proportionality) to be used for all admissible profiles.

As discussed in Section 2, any increasing transform of a von Neumann–Morgenstern utility function represents the same preference. As Sen has noted, if we subject the individual utility functions in \mathbf{U} to non-affine transforms, then we can represent the social preference relation by a nonlinear function of the individual utilities. Hence, the social objective function appears to be weighted utilitarian with one representation of the individual preferences and non-utilitarian with another.¹⁶

In order to illustrate this objection to Harsanyi’s utilitarian interpretation of his Aggregation Theorem, I employ a variant of the example used in Weymark (1991, Section 4). Let \mathbf{V} be the profile that is obtained by subjecting each of the individual utility functions in \mathbf{U} to an exponential transform. Thus, $V_i = \exp(U_i)$ for $i = 1, \dots, n$. Recall that each of the functions U_i are von Neumann–Morgenstern utility functions whose images are nondegenerate intervals. Thus, none of the V_i are von Neumann–Morgenstern representations. Nevertheless, both U_i and V_i represent \succeq_i . For this reason, or so Sen would argue, \mathbf{V} is an admissible profile of utility functions. Because the ex-

¹⁶See, for example, Sen (1986, pp. 1123–1124) where this point is made in the context of Harsanyi’s Impartial Observer Theorem. Harsanyi (1955, 1977b) regards his Aggregation and Impartial Observer Theorems, as well as a third theorem due to Fleming (1952), as providing independent justifications for utilitarianism. However, Blackorby, Donaldson, and Weymark (1980) have argued that Harsanyi has offered no argument as to why the same representations of individual preferences must be used in each of these theorems, and so has failed to establish that they even lead to the same social rankings of the alternatives, let alone to some form of utilitarianism.

ponential function is increasing, it can be inverted. Thus, $U_i = \ln(V_i)$ for $i = 1, \dots, n$. Substituting these expressions into (5), we obtain

$$\begin{aligned} p \succeq q &\leftrightarrow \sum_{i=1}^n a_i \ln(V_i(p)) \geq \sum_{i=1}^n a_i \ln(V_i(q)) \\ &\leftrightarrow \prod_{i=1}^n V_i(p)^{a_i} \geq \prod_{i=1}^n V_i(q)^{a_i}, \quad \forall p, q \in \mathcal{L}. \end{aligned} \tag{7}$$

Therefore, using the utility representations in \mathbf{U} , it appears that the social ranking \succeq of \mathcal{L} is obtained using a weighted sum of the individual utilities, whereas using the utility representations in \mathbf{V} , it appears that \succeq is obtained using a Cobb–Douglas function of these utilities.

The underlying reason for the problems identified by Sen is that in order for utilitarianism, either in its weighted or classical form, to be meaningful, it must be possible to compare utility differences (gains and losses) both intrapersonally and interpersonally. However, if the only information about individuals that is available is their orderings of the set of alternatives, it is not possible to make such comparisons. The need for difference comparability can be seen most clearly by rewriting (5) as

$$p \succeq q \leftrightarrow \sum_{i=1}^n a_i [U_i(p) - U_i(q)] \geq 0, \quad \forall p, q \in \mathcal{L}. \tag{8}$$

The sum in (8) is not, in general, invariant to independent increasing transforms of the individual utility functions, even if these transforms are restricted to be affine.

In order for (8) to hold for all $\mathbf{U} \in \mathcal{U}^a$, the transforms that are applied to the individual utility functions cannot be chosen independently. Suppose, as above, that the weights in (5), and hence in (8), are obtained from Theorem 2 using the profile \mathbf{U} of von Neumann–Morgenstern utility functions and that Independent Prospects is satisfied. Consider the n -tuple of transforms $F = (f_1, \dots, f_n)$, where for all $i = 1, \dots, n$, $f_i(t) = a_i + bt$ for all $t \in \mathbb{R}$ for some real numbers (a_1, \dots, a_n, b) with $b > 0$. Such an n -tuple of transforms is called *co-cardinal*. It is clear that the utility sum in (8) does not change if the utility profile \mathbf{U} is replaced by the profile $\mathbf{V} = F \circ \mathbf{U} := (f_1 \circ U_1, \dots, f_n \circ U_n)$ for some co-cardinal n -tuple of transforms F . Let \mathcal{U}^c denote the set of such profiles of utility functions. Because the images of the utility functions in \mathbf{U} are nondegenerate intervals, the profiles in \mathcal{U}^c are the only profiles that

preserve the utility sum in (8).¹⁷ Hence, if the set of admissible profiles of utility functions \mathcal{U}^a is a subset of \mathcal{U}^c , then it is legitimate to say that \succeq is a weighted utilitarian social preference.

In Harsanyi (1977b, p. 293) and Harsanyi (1979, pp. 296–297), an argument is presented that appears to demonstrate that if an individual’s preferences conform to the axioms of expected utility theory, then it is meaningful to make intrapersonal comparisons of utility differences. As discussed above, the possibility of making such comparisons is a necessary (but not sufficient) condition for utilitarianism to be meaningful. It is therefore instructive to sketch this argument.

Suppose that individual i is asked to rank the lotteries $A = [0.5p + 0.5q]$ and $B = [0.5r + 0.5s]$, where $p, q, r, s \in \mathcal{L}$. Without loss of generality, suppose that $A \succeq_i B$. If \succeq_i is represented by the von Neumann–Morgenstern utility function U_i , it follows from the Expected Utility Theorem that

$$0.5U_i(p) + 0.5U_i(q) \geq 0.5U_i(r) + 0.5U_i(s) \quad (9)$$

or, equivalently, that

$$U_i(p) - U_i(s) \geq U_i(r) - U_i(q). \quad (10)$$

Furthermore, the inequality in (10) is preserved if U_i is subjected to any increasing affine transform. Suppose that only von Neumann–Morgenstern utility functions are legitimate representations of a preference that satisfies the expected utility theory axioms discussed in Section 2. If this were true, it would then follow from Theorem 1 that it is meaningful to make intrapersonal comparisons of utility differences.¹⁸ But, as we have seen, nothing in the version of expected utility theory that Harsanyi employed in his theorems

¹⁷In the absence of the nondegeneracy assumption, the set of profiles of utility functions in \mathcal{U}^c is a strict subset of the set of profiles of utility functions that preserve the utility difference comparisons exhibited by \mathbf{U} . See Ellingsen (1994) or Bossert and Weymark (2004).

¹⁸In a letter to the author dated October 29th, 1990, Harsanyi has informed me that the choice of a von Neumann–Morgenstern representation is only required if “we want a utility function representing both the ordinal and the cardinal properties of his preferences and of his satisfactions.” As best as I can determine from this letter, Harsanyi is supposing that the individual in question also has a preference over pairs of alternatives, which permits him to make meaningful statements about the strength of preference for, say, A over B relative to C over D . If this is correct, Harsanyi’s views are similar to those of Broome (1991) discussed in Section 7.1 below.

rules out the use of a non-affine, increasing transform of U_i . Provided that none of the lotteries p, q, r , and s are indifferent to each other, it is always possible to find such a transform for which the inequality in (10) does not hold. Thus, it appears that expected utility theory does not permit one to make intrapersonal comparisons of utility differences.¹⁹

Recall that an impartial observer considers lotteries in which not only the physical outcome, but also his identity, is uncertain. Provided that the impartial observer's preferences over such extended lotteries satisfy the expected utility axioms and provided that only von Neumann–Morgenstern representations of these preferences are admissible, the argument presented in the preceding paragraph can be extended to show that interpersonal utility comparisons are meaningful. The details of this argument may be found in Weymark (1991, p. 306). However, as above, it does not follow from the version of expected utility theory considered in Section 2 that only von Neumann–Morgenstern utility representations of the impartial observer's preferences are permissible, and so it is illegitimate to use this argument to conclude that interpersonal comparisons of utility differences are meaningful.

It should be stressed that nothing in the preceding discussion permits one to conclude that it is impossible to make interpersonal comparisons of utility differences. Indeed, Harsanyi has made fundamental contributions to our understanding of the logical basis for making interpersonal utility comparisons. See, for example, Harsanyi (1955, 1977b). Harsanyi clearly thinks of utility as being more than just a representation of a preference and there is considerable textual evidence to suggest that Harsanyi's arguments in support of utilitarianism are based, in part, on a non-representational concept of utility, even though it is not employed in his formal theorems.²⁰

5. Measurement Theory

In their discussions of the cardinality of von Neumann–Morgenstern utility functions, suitably reinterpreted as representations of well-being, Broome (1991, 1997) and Risse (2002) have either explicitly or implicitly made use of results in formal measurement theory. In order to evaluate their arguments and their relevance for utilitarianism, it is first necessary to review the

¹⁹The belief that such comparisons are meaningful was so widespread that Luce and Raiffa (1957, p. 32) cite it as being one of the common fallacies of expected utility theory.

²⁰See Weymark (1991, Section 6) for a detailed discussion of this issue.

relevant parts of measurement theory.

The early analyses of the nature of measurement, such as that of Campbell (1920), were concerned with measurement in the physical sciences. While there were important contributions to the theory of measurement in the social sciences prior to the mid-1950s, of which von Neumann and Morgenstern’s axiomatization of expected utility theory is a prominent example, it was not until the seminal work of Scott and Suppes (1958), Pfanzagl (1959a), and Suppes and Zinnes (1963) that the modern representational theory of measurement was first systematized. This approach to measurement is broad enough to encompass the measurement issues that arise in both the physical and social sciences. Good introductions to measurement theory may be found in Krantz, Luce, Suppes, and Tversky (1971) and Roberts (1979).²¹

In measurement theory, a *relational structure* is a set S together with one or more relations on S . There are two kinds of relational structures, empirical and numerical. In an *empirical relational structure*, the set S is the set of objects to be measured, whereas in a *numerical relational structure*, the set S is a set of numbers. A measurement involves constructing a mapping h , called a *homomorphism*, between an empirical relational structure \mathfrak{E} and a numerical relational structure \mathfrak{N} that preserves all of the relations in \mathfrak{E} . The triple $(\mathfrak{E}, \mathfrak{N}, h)$ is called a *scale*.²² This formalization of measurement is general enough to encompass binary operators on $S \times S$ because such an operator can be equivalently thought of as being a ternary relation on S .²³

These concepts can be illustrated with the measurement of weight. The empirical relational structure is $\mathfrak{W} = \langle S; \succeq; \circ \rangle$, where S is the set of objects to be weighed, \succeq is a binary relation on S that is interpreted as “weighs at least as much as,” and \circ is a concatenation operator that combines the ordered pair of objects $(x, y) \in S \times S$ to produce a new object $x \circ y \in S$.²⁴ The relational

²¹There is some dispute about what phenomena are quantitatively measurable. Berka (1983, p. 63), for one, argues that quantitative measurement presupposes that it is possible to perform “operations of measurement with the help of standard measuring instruments which must be *constant, easily reproducible* and suitably *precise* (emphasis in the original).” He questions whether this is possible for extraphysical phenomena such as utility. See Berka (1983, Sections 7.2 and 8.2) for an elaboration of this argument that specifically considers von Neumann–Morgenstern utility theory.

²²For given relational structures \mathfrak{E} and \mathfrak{N} , it is common to simply refer to h as the scale. When there is no ambiguity, I shall follow this practice in the subsequent discussion.

²³See Krantz, Luce, Suppes, and Tversky (1971, p. 8).

²⁴When considering binary operators, it is always assumed that the set S is closed under this operation; i.e., that the object obtained by applying the operator to any $(x, y) \in S \times S$

structure \mathfrak{W} is the formal expression of the empirical procedure used to weigh objects using a two-pan balancing scale. If object x is put in one pan and object y in the other, then $x \succeq y$ if the height of the pan containing x does not exceed the height of the pan containing y . The concatenation $x \circ y$ corresponds to putting both x and y in the same pan of the scale, say x on top of y .

The usual practice in the theory of weight is to associate the numerical structure $\mathfrak{N}^1 = \langle \mathbb{R}_+; \geq; + \rangle$ with \mathfrak{W} . For each $x \in S$, the homomorphism $w: \mathfrak{W} \rightarrow \mathfrak{N}^1$ assigns the weight $w(x) \geq 0$ to x . The ordering \succeq is preserved by requiring that

$$w(x) \geq w(y) \leftrightarrow x \succeq y, \quad \forall x, y \in S, \quad (11)$$

and the concatenation operator \circ is preserved by requiring that

$$w(x \circ y) = w(x) + w(y), \quad \forall (x, y) \in S \times S. \quad (12)$$

In other words, the numerical relation “greater than or equal to” corresponds to the empirical relation “weighs at least as much as” and the weight of a pair of objects is simply the sum of their individual weights.

The standard way of measuring length has the same formal structure as that of measuring weight, so we simply need to reinterpret the components of $(\mathfrak{W}, \mathfrak{N}^1, w)$ in order to obtain a measurement scale for length. In this new interpretation, S is the set of one-dimensional objects (rods) whose lengths are to be determined, \succeq is interpreted as “is at least as long as,” and \circ is again a concatenation operator that combines the ordered pair of objects (x, y) to produce a new object $x \circ y$. One possible physical interpretation of \succeq is that $x \succeq y$ if the rods x and y are aligned side by side with one end of x adjacent to one end of y and the other end of y does not extend past the other end of x . The operator \circ could correspond to placing two rods end to end. The function w now assigns a length to each rod in such a way that the ordering of the lengths is preserved and such that the length of a concatenated rod is the sum of the lengths of the two rods that were used to construct it.

Two basic problems that a satisfactory theory of measurement must solve are the representation and uniqueness problems. Given an empirical relational structure \mathfrak{E} and a numerical relational structure \mathfrak{N} , the *representation*

is also in S .

problem is to find sufficient (and, ideally, necessary) conditions on \mathfrak{E} for the existence of a homomorphism h from \mathfrak{E} to \mathfrak{N} that preserves the relations in \mathfrak{E} . For example, in the case of the measurement of weight or length, the requisite homomorphism exists if \mathfrak{W} is what is known as a closed extensive structure when the numerical relational structure is \mathfrak{N}^1 .²⁵

The *uniqueness problem* is concerned with determining the class of transformations that when applied to a homomorphism h that solves the representation problem (given \mathfrak{E} and \mathfrak{N}) yields another homomorphism that also solves this problem. A uniqueness theorem identifies the kind of scale h is (i.e., the group structure of the class of transformations that may be applied to h).²⁶ For example, in the case of weight or length, the homomorphism is unique up to a similarity transform. A function $f: \mathbb{R} \rightarrow \mathbb{R}$ is a *similarity transform* if there exists a real number $a > 0$ such that $f(t) = at$ for all $t \in \mathbb{R}$. Representations that are unique up to a similarity transform are called *ratio scales*.²⁷

There is a further problem that needs to be addressed that has received much less attention in the literature. Given the empirical relational structure \mathfrak{E} , on what basis should we choose the numerical relational structure \mathfrak{N} that \mathfrak{E} is to be mapped into? Roberts (1979, p. 54) refers to this issue as being a “difficult philosophical question—not a mathematical question.”

The nature of this problem can be illustrated by reconsidering the measurement of weight and length. Following standard practice, in the preceding discussion, I have mapped the empirical relational structure \mathfrak{W} into the numerical relational structure \mathfrak{N}^1 . Another possibility, which is considered by Krantz, Luce, Suppes, and Tversky (1971, Sections 1.2.3 and 3.9), is to replace \mathfrak{N}^1 with $\mathfrak{N}^2 = \langle \mathbb{R}_+; \geq; \times \rangle$ and instead of requiring the homomorphism w to satisfy (12), require that

$$w(x \circ y) = w(x) \times w(y), \quad \forall (x, y) \in S \times S. \quad (13)$$

In other words, the weight (or length) of the concatenated object $x \circ y$ is given by the product of the weights (or lengths) of x and y . Thus, instead of

²⁵See Krantz, Luce, Suppes, and Tversky (1971, p. 73) for the definition of this structure.

²⁶More precisely, the set of admissible transformations together with the function composition operator form an algebraic group. See Bossert and Weymark (2004) for a detailed discussion of groups of transformations.

²⁷The importance of identifying the scale type in a measurement exercise was stressed by Stevens (1946), who also identified many of the scale types commonly encountered in measurement theory.

seeking an additive representation of \mathfrak{W} , we require the representation to be multiplicative.

Note that if $w: S \rightarrow \mathbb{R}_+$ satisfies (11) and (12), then $\bar{w}: S \rightarrow \mathbb{R}_+$ defined by setting $\bar{w}(x) = \exp(w(x))$ for all $x \in S$ satisfies (11) and (13). As is the case with the additive representation w , the multiplicative representation \bar{w} is unique up to a one-parameter class of transforms. Either w or \bar{w} could serve as a measure of weight, length, or any other closed extensive structure. Furthermore, \mathfrak{N}^1 and \mathfrak{N}^2 are just two of the infinite number of possible choices for the numerical relational structure. In other words,

... despite its great appeal and universal acceptance, the additive representation is just one of the infinitely many, equally adequate representations [of an extensive structure] The essential fact about the uniqueness of the representation is not the particular group of admissible transformations, but that all groups are isomorphic and, in the case of extensive measurement, are all one-parameter groups (Krantz, Luce, Suppes, and Tversky, 1971, p. 102)

The possible appeal of a non-additive representation for a closed extensive structure, such as weight or length, can be seen by considering an alternative empirical interpretation of the concatenation operation used in the measurement of length to the one considered above. Consider the empirical relational structure $\mathfrak{L} = \langle S; \succeq; * \rangle$, where S and \succeq have the same physical interpretations as in our earlier discussion of length. The object $x * y$ obtained by concatenating x and y is now interpreted to be the hypotenuse of the right-angle triangle obtained by orthogonally abutting y to one end of x .²⁸ It is not difficult to verify that \mathfrak{L} is a closed extensive structure, so there exists an additive representation l of “orthogonal” length using the numerical relational structure \mathfrak{N}^1 . Using Pythagoras’ Theorem, it follows that such a measure can be obtained from an additive measure w obtained using the usual concatenation operator \circ by letting $l = w^2$. Suppose, however, that we want length to be additive in terms of our original concatenation operator \circ , rather than additive in terms of $*$. In that case, as has been shown by Krantz, Luce, Suppes, and Tversky (1971, p. 99), we should use the numerical relational

²⁸This interpretation of the empirical structure used for measuring length was initially proposed by Ellis (1966). See also the discussion of this example in Berka (1983, pp. 155–157) and Krantz, Luce, Suppes, and Tversky (1971, Sections 3.6.1 and 3.9).

structure $\mathfrak{N}^3 = \langle \mathbb{R}_+, \geq, \oplus \rangle$, where

$$x \oplus y = x + y + 2\sqrt{xy}, \quad \forall(x, y) \in \mathbb{R}_+^2, \quad (14)$$

and this results in a non-additive measure of length in terms of the operator $*$.²⁹

In practice, it has been pragmatic considerations that have dictated the choice of the numerical relational structure used in a measurement exercise. For example, in the case of weight or length, the computations that are required to compute the weight or height of a concatenated object are simpler, or at least more familiar to us, when an additive representation is used than when, say, the representation is multiplicative. Furthermore, the mathematical form of the physical laws that employ weight or length may be simpler with an additive representation than with an alternative form.³⁰ Nevertheless,

[t]his choice is essentially a matter of convention, although the conventions are strongly affected by computational convenience.
(Krantz, Luce, Suppes, and Tversky, 1971, p. 102)

6. Expected Utility Theory Reconsidered

As their extensive discussion of the principles of quantitative measurement and their discussion of measurement in the physical sciences make clear, von Neumann and Morgenstern (1944, Chapter 1, Section 3) regard themselves as providing a theory of measurable utility in the same sense that heat, length, and other physical phenomena are measurable. Von Neumann and Morgenstern are quite rightly acknowledged as being pioneers in the extension of measurement from the physical to the social sciences.³¹ However, von

²⁹Berka (1983, pp. 156) argues that it is a mistake to think that abutting two rods end to end and abutting them orthogonally are two different interpretations of the concatenation operation for measuring length. He regards the two operations as relating to the measurement of different “magnitudes” and suggests that an additive representation is appropriate in either case. However, the fact that \mathfrak{M} and \mathfrak{L} are different empirical relational structures is not at issue here. Rather, it is what numerical relational structures should be associated with them.

³⁰See Ellis (1966, pp. 81–86) for a defence of the view that the choice of a representation should be determined by the simplicity of the resulting physical laws or computations.

³¹See, for example, Pfanzagl (1959b), who uses von Neumann–Morgenstern utility theory to illustrate the then newly developed representational theory of measurement.

Neumann and Morgenstern’s presentation of their theory is not easy to understand. In part, this is because they used “utility” both to describe the set of objects that are being measured and as the numbers that are used in their representation theorem.³² It is for this reason that the subsequent literature has abandoned von Neumann and Morgenstern’s formalization of expected utility theory in favour of models based on the reformulations of the theory found in Marschak (1950) and Harsanyi and Milnor (1953).³³ However, in these latter expositions of expected utility theory, the measurement-theoretic role that binary operators play in the theory is less prominent than it is in von Neumann and Morgenstern (1944).

In order to determine whether von Neumann–Morgenstern utility functions are cardinal in the sense needed for Harsanyi’s theorems, it is necessary to reconsider von Neumann–Morgenstern utility theory in terms of the representational theory of measurement developed in the preceding section. As in Section 2, I employ Marschak’s version of expected utility theory in my discussion. The same conclusions can be obtained using the original formalization of von Neumann and Morgenstern, as I shall briefly argue at the end of this section.

As described in the preceding section, measurement involves specifying empirical and numerical relational structures and a homomorphism between them. In Marschak’s version of expected utility theory, the empirical relational structure is $\mathfrak{M} = \langle \mathcal{L}; \succeq; \alpha_k, k \in [0, 1] \rangle$, where, as in Section 2, \mathcal{L} is a set of lotteries on a set of sure alternatives X and \succeq is a weak preference binary relation on \mathcal{L} . For all $k \in [0, 1]$, α_k is the binary operator defined by setting

$$p\alpha_kq = kp + (1 - k)q, \quad \forall(p, q) \in \mathcal{L} \times \mathcal{L}. \quad (15)$$

Note that there are an infinite number of these operators, one for each number in $[0, 1]$. Thus, taking convex combinations of lotteries are the binary operators in the empirical relational structure for Marschak’s version of expected utility theory. It is these operators that are the counterparts to the concatenation operators that appear in the measurement of closed extensive structures such as weight and length. The standard numerical relational

³²In their informal discussion, von Neumann and Morgenstern also consider the lottery formulation of expected utility theory used here. For a very useful exegesis of the sections on expected utility theory in von Neumann and Morgenstern (1944), see Fishburn (1989).

³³Harsanyi and Milnor (1953) develop their axiomatization of expected utility theory for general mixture sets. The set of lotteries \mathcal{L} is a mixture set.

structure for expected utility theory is $\mathfrak{N}^4 = \langle \mathbb{R}; \succeq; \bar{\alpha}_k, k \in [0, 1] \rangle$, where for all $k \in [0, 1]$, $\bar{\alpha}_k$ is the numerical binary operator defined by setting

$$x\bar{\alpha}_ky = kx + (1 - k)y, \quad \forall(x, y) \in \mathbb{R}^2. \quad (16)$$

The homomorphism relating \mathfrak{M} to \mathfrak{N}^4 is a function $U: \mathcal{L} \rightarrow \mathbb{R}$ that is order preserving and that preserves the linear structure of the convex combination operators in \mathfrak{M} . The first condition simply requires that U be a utility function; i.e., that U satisfies (1). The second condition requires that for all $k \in [0, 1]$,

$$U(p\alpha_kq) = U(p)\bar{\alpha}_kU(q), \quad \forall(p, q) \in \mathcal{L} \times \mathcal{L}, \quad (17)$$

or, equivalently, that

$$U(kp + (1 - k)q) = kU(p) + (1 - k)U(q), \quad \forall(p, q) \in \mathcal{L} \times \mathcal{L}. \quad (18)$$

It is not difficult to verify that (17) and (18) are both equivalent to (2). Hence, U must be a von Neumann–Morgenstern utility function. From the Expected Utility Theorem (Theorem 1), we know that such a homomorphism exists if \succeq satisfies Ordering, Continuity, and Independence. Furthermore, if it is additionally assumed that Nondegeneracy is satisfied, then U is unique up to an increasing affine transform.

It thus appears that by taking explicit note of the binary operators α_k that expected utility theory is cardinal. Indeed, according to the precepts of representational measurement theory, it is cardinal for, given the specification of the two relational structures \mathfrak{M} and \mathfrak{N}^4 , the admissible homomorphisms are unique up to an increasing affine transform. Nevertheless, this observation does not allow us to conclude that we must use a von Neumann–Morgenstern representation of the preference \succeq . This conclusion only holds if we assume that the numerical relational structure is \mathfrak{N}^4 .

Suppose instead that we use the numerical relational structure $\mathfrak{N}^5 = \langle \mathbb{R}; \succeq; \hat{\alpha}_k, k \in [0, 1] \rangle$, where for all $k \in [0, 1]$, $\hat{\alpha}_k$ is the ‘‘Cobb–Douglas’’ operator defined by setting

$$x\hat{\alpha}_ky = x^ky^{(1-k)}, \quad \forall(x, y) \in \mathbb{R}^2. \quad (19)$$

With this numerical relational structure, in order to preserve the binary operators in \mathfrak{M} , for all $k \in [0, 1]$, the utility representation U is now required to satisfy

$$U(kp + (1 - k)q) = U(p)^kU(q)^{(1-k)}, \quad \forall(p, q) \in \mathcal{L} \times \mathcal{L}, \quad (20)$$

rather than (18).

From the discussion in Section 4, it is clear that any utility representation of \succeq that satisfies (20) is an exponential transform of a utility representation of \succsim that satisfies (18). Thus, if Marschak's four axioms are satisfied, the class of admissible representations for the numerical relational structure \mathfrak{N}^5 is also characterized by a two-parameter family of transformations. It follows from the arguments in Weymark (1991, pp. 283–284) that a transform $f: \mathbb{R}_+ \rightarrow \mathbb{R}$ is in this family if and only if there exist $a, b \in \mathbb{R}$ with $b > 0$ such that

$$f(t) = \exp[a + b \ln(t)], \quad \forall t \in \mathbb{R}_+. \quad (21)$$

Consider two representations U and V of \succeq for which $V = \exp[a + b \ln(U)]$, where $b > 0$. Note that for all $p, q, r, s \in \mathcal{L}$,

$$\begin{aligned} \frac{V(p)}{V(q)} &\geq \frac{V(r)}{V(s)} \\ &\Leftrightarrow \\ \ln(V(p)) - \ln(V(q)) &\geq \ln(V(r)) - \ln(V(s)) \\ &\Leftrightarrow \\ \ln(U(p)) - \ln(U(q)) &\geq \ln(U(r)) - \ln(U(s)) \\ &\Leftrightarrow \\ \frac{U(p)}{U(q)} &\geq \frac{U(r)}{U(s)}. \end{aligned} \quad (22)$$

Thus, the class of transforms in (21) preserves comparison of ratios of utilities, not utility differences.

As with weight and length, the choice of \mathfrak{N}^4 rather than \mathfrak{N}^5 for the numerical relational structure in expected utility theory has been dictated by pragmatic considerations of the kind described in Section 2 for choosing von Neumann–Morgenstern representations from among all the possible representations of \succeq . This is perfectly acceptable when describing choice behaviour, which is the objective von Neumann and Morgenstern (1944, p. 20) offer for their theory of measurable utility. If one is only interested in describing or predicting choice behaviour, it doesn't matter which numerical relational structure is used (provided that it is isomorphic to \mathfrak{N}^4), and so pragmatic considerations may be decisive. However, this is not a satisfactory basis for

choosing a numerical relational structure for the measurement of utility when the objective is to provide a normative foundation for utilitarianism.

As noted above, von Neumann and Morgenstern (1944) did not use \mathfrak{M} as their empirical relational structure in their formal theory. Instead, they used $\mathfrak{U} = \langle \mathcal{V}; \succ; \alpha_k^*, k \in [0, 1] \rangle$, where \mathcal{V} is an “abstract” set of utilities, \succ is a strict preference relation on \mathcal{V} , and the α_k^* are operators that form “convex combinations” of ordered pairs of abstract utilities. Their numerical relational structure is $\mathfrak{N}^6 = \langle \mathbb{R}; >; \hat{\alpha}_k, k \in [0, 1] \rangle$, which is simply \mathfrak{N}^4 with the strict inequality relation $>$ on \mathbb{R} substituting for \geq .³⁴ Fishburn (1989, p. 138) suggests interpreting \mathcal{V} as “an abstract set of indifference classes”. The structures \mathfrak{M} and \mathfrak{U} have similar formal properties, so it is possible to reformulate the preceding argument using von Neumann and Morgenstern’s version of expected utility theory.

In terms of a contribution to measurement theory, von Neumann and Morgenstern (1944, p. 24) clearly regard their discovery of a “natural” operation which can be used to restrict the class of utility transformations found in ordinal utility theory to the class of affine transformations that appear in their representation theorem as constituting their greatest achievement. Here, von Neumann and Morgenstern are referring to the operators that appear in their empirical relational structure; i.e., to the convex combination operators. These are operators that are concerned with the “natural” world, not the world of real numbers. For von Neumann and Morgenstern, and for most of the scholars who have followed in their footsteps, it seems obvious that the convex combination operator $\bar{\alpha}_k$ for real numbers should be chosen to correspond to the empirical operator α_k^* (or α_k). According to von Neumann and Morgenstern (1944, p. 24), these empirical and numerical operators are “synonymous.” The choice of $\bar{\alpha}_k$ as the numerical operator has considerable appeal because it has many formal properties in common with α_k^* and α_k . It was only as the theory of measurement was further developed that this choice was seen as needing further justification.

7. Broome and Risse on the Cardinality of Utility

Broome (1991, 1997) and Risse (2002) have appealed to measurement theory to provide support for the claim that Harsanyi’s theorems have some rele-

³⁴See Krantz, Luce, Suppes, and Tversky (1971, p. 407) for a succinct formalization of von Neumann and Morgenstern’s model.

vance for utilitarianism. In spite of some similarities between Broome’s arguments in his monograph and the latter arguments in Broome’s and Risse’s articles, in fact these arguments are fundamentally quite different.³⁵ Their analyses consider many issues related to the significance of Harsanyi’s theorem for utilitarianism. It is beyond the scope of this article to consider all of these issues here. Rather, I shall focus my attention on their claims that their concepts of *individual* well-being or goodness are cardinal in the sense needed for Harsanyi’s theorems to be interpreted as being theorems about utilitarianism, at least in its weighted utilitarian formulation.

It is my contention that Broome (1991) has provided an account of utility in which intrapersonal comparisons of utility differences are meaningful, but he has only done this by introducing an additional quaternary operator that is not part of the von Neumann–Morgenstern approach to expected utility theory. On the other hand, Broome (1997) and Risse (2002) suffer from the problems related to the choice of numerical relational structure raised in the preceding section. Broome recognizes that the choice of this structure requires justification, but does not offer a normatively compelling reason for his choice of the standard structure \mathfrak{N}^4 . Risse does not comment on this issue.

Of course, in order for Harsanyi’s theorems to provide support for utilitarianism, differences in utility must be comparable interpersonally, not just intrapersonally. However, if, as I believe the preceding section has demonstrated, von Neumann–Morgenstern utility theory does not supply a normatively compelling cardinal measure of individual utility, then Harsanyi has not demonstrated what he set out to achieve. Nevertheless, if one is convinced by the arguments in Broome (1991) and is willing to go beyond the von Neumann–Morgenstern theory by supposing that differences in utility are quantitatively measurable, then a case can be made for the relevance of Harsanyi’s theorems for utilitarianism.

7.1. Broome (1991)

For Broome (1991), utility is a measure of the goodness of alternatives and goodness in turn is described in terms of a betterness binary relation (“weakly better than”) on these alternatives. In Broome’s formulation of Harsanyi’s

³⁵Indeed, as we shall see, Risse explicitly appeals to an argument made by Broome (1991) when presenting his case for the cardinality of well-being.

Aggregation Theorem, individual and social betterness relations are used instead of individual and social preference relations. Broome employs a state-contingent alternatives version of expected utility theory.³⁶ There are S states of nature, $s = 1, \dots, S$. Each of these states occurs with some fixed positive probability. An *act* (or prospect) a is a mapping that assigns a sure outcome from some set X to each state. Let A denote the set of acts. For each $x_m \in X$, A includes the constant act in which x_m is obtained in every state.

My concern is with what can be said about the measurability of goodness as described by an individual betterness relation \succeq_b on A . While expressing some reservations, Broome argues that it is reasonable to require this relation to conform to the expected utility axioms for the state-contingent model with fixed probabilities. If this is the case, then \succeq_b can be represented by an expectational utility function; i.e., a function for which the utility of an act is the expected utility of its state-contingent outcomes, where the utility of an outcome $x_m \in X$ is simply the utility assigned to the constant act associated with x_m . Such an expectational representation is unique up to an increasing affine transform. However, as Broome recognizes, any increasing transform of an expectational representation also represents the betterness relation \succeq_b . The claim that only expectational representations measure goodness is what Broome (1991, p. 142) calls *Bernoulli's hypothesis*. It is to provide support for this hypothesis that Broome appeals to measurement theory, albeit rather informally.

The title, *Weighing Goods*, of his monograph suggests that Broome draws an analogy between the measurement of weight and the measurement of goodness. However, the sense in which he uses “weighing” is very different from that used in Section 5 when discussing weight measurement. For Broome, it appears that “weighing” is best thought of as referring to weighing reasons for and against different choices. This difference is not immediately apparent because Broome never provides a complete mathematical description of the model he employs, relying instead on verbal arguments illustrated with numerical examples. By analyzing the example Broome (1991, pp. 146–148) uses to defend Bernoulli’s hypothesis, we can see why I believe that his argument depends on the use of a second relation to supplement the information provided by \succeq_b .

³⁶See Blackorby, Donaldson, and Weymark (1999) for a formal statement and proof of Harsanyi’s Aggregation Theorem for this model of uncertainty.

In Broome’s example, there are two equally-likely states of nature, H and T (e.g., the outcomes obtained by flipping a fair coin). In prospect a^1 , the individual receives £100 if state H occurs and £200 if state T occurs, whereas in prospect a^2 the outcomes are £20 and £320, respectively. Determining which prospect is better is for Broome (1991, p. 146)

... a matter of weighing against each other *differences in good* located in different states of nature. The consideration in favour of prospect $[a^1]$ is that if state H occurs, it gives you the good of £100 rather than the good of £20. The consideration in favour of $[a^2]$ is that if state T occurs, it gives you the good of £320 rather than the good of £200. These two differences—each a *difference in good*—have to be weighed against each other. (my emphasis)

In the case in which these differences exactly balance each other, the utility representation must equate the utility difference in going from £20 to £100 with the utility difference in going from £200 to £320. Broome (1991, pp. 147) continues by saying that:

The fact that the two utility differences are the same tells us, from the definition of utility, that the consideration in favour of $[a^1]$ exactly balances the consideration in favour of $[a^2]$. Utility, that it to say, tells us how much the differences in good *count* in determining the overall goodness for you of the alternative prospects.

Since the two differences in good are exactly balanced in determining the overall goodness of the prospects, it would be very natural to express this fact by saying that the differences are actually the same. ... So to deny they are actually the same would be to insist on a distinction between amounts of good and how much those amounts count in determining overall goodness. And it is natural to think this an empty distinction. (emphasis in the original)³⁷

If I have interpreted Broome correctly, he has supplemented the binary relation \succeq_b with a quaternary relation \succeq_d on X . The statement that $(x_1, x_2) \succeq_d (x_3, x_4)$ is interpreted as saying that the difference in goodness obtained by

³⁷Note that Broome’s use of “natural” in this quotation is different in meaning from the use of “natural” in von Neumann and Morgenstern (1944).

replacing x_2 with x_1 is at least as great as the difference in goodness obtained by replacing x_4 with x_3 .³⁸ It is this “difference in goodness” relation that provides the cardinality that Broome needs.³⁹ In effect, Broome has combined the von Neumann and Morgenstern (1944) approach to expected utility with a theory based on directly comparing differences in the goodness of alternatives, as in Pareto (1906).

Note that it is possible to restate this argument in terms of the lotteries \mathcal{L} . Simply let $X = \{\mathcal{L}20, \mathcal{L}100, \mathcal{L}200, \mathcal{L}320\}$ and let a^1 (resp. a^2) be the lottery that assigns 50% probability to both $\mathcal{L}100$ and $\mathcal{L}200$ (resp. $\mathcal{L}20$ and $\mathcal{L}320$). With this reinterpretation, everything Broome says about the state-contingent model applies equally well to the lottery model.

7.2. Risse (2002)

As in Broome (1991), Risse (2002) is concerned with a betterness relation that satisfies the axioms of expected utility theory. In his case, the betterness relation is entailed by an idealized desire satisfaction concept of well-being and the set of alternatives is the set of lotteries \mathcal{L} . Risse provides a clear and accurate account of von Neumann and Morgenstern’s contribution to measurement theory. In particular, he recognizes the importance of the convex combination operators in von Neumann and Morgenstern’s axiomatization of expected utility theory. Risse (2002, p. 563) says that:

The additional natural operation (i.e., in addition to “preferring”) that von Neumann and Morgenstern think they discovered is the *concatenation of events with probabilities*. . . . And then it only takes the axiomatic postulation of properties of the concatenation operation to obtain a measurement of utility based on these two natural operations. Those properties must be chosen so that the behavior of numerical utilities captures the expectational nature of utility.⁴⁰ (emphasis in the original)

³⁸See also Broome (1993, Section 2), where it is said that “betterness” is a cardinal concept, and cardinality is defined in terms of the meaningfulness of difference comparisons.

³⁹For consistency, the ranking of certain acts by \succeq_b must coincide with the ranking of the outcomes in X that is implicit in \succeq_d . See Krantz, Luce, Suppes, and Tversky (1971, Chapter 4) and Roberts (1979, Chapter 3) for discussions of difference measurability that employ the framework of formal measurement theory.

⁴⁰An “event” is what von Neumann and Morgenstern (1944) call a sure outcome.

At this point in the discussion, Risse’s argument becomes a bit unclear. If I have understood him correctly, he does not recognize that von Neumann and Morgenstern’s choice of a numerical relational structure needs justification, and so accepts the conclusion that von Neumann and Morgenstern have determined

... a family of functions closed under positive affine transformations that measure utility just as positive affine transformations of the Celsius scale measure heat. The family of functions, of course, is the family of expectational representations. (Risse (2002, pp. 563–564))

For (Risse, 2002, p. 564), what remains to be established is “that the family of expectational representations does indeed measure well-being.” In other words, a representation of the betterness relation measures well-being (not just utility) if and only if it is a von Neumann–Morgenstern utility function. To complete this final step in his argument, Risse (2002, pp. 564–565) appeals to Broome’s argument described above, but now applied to an idealized desire satisfaction interpretation of well-being, rather than goodness. However, as I shall explain, it seems to me that Risse’s argument differs from Broome’s in a fundamental way.

Although in the rest of his article, alternatives are lotteries, for simplicity, Risse switches to state-contingent alternatives when appealing to Broome’s argument, fully recognizing that this argument can be restated in terms of lotteries. He uses the same example as Broome, but with the probabilities of the two states O_1 (H) and O_2 (T) now $1/3$ and $2/3$, respectively, and dollars substituted for pounds. He assumes that the two prospects are indifferent to each other according to the betterness relation.

Risse (2002, pp. 564) says that:

Since we are assuming that the betterness-relation entailed by your well-being can be represented by an expectational function u , we obtain the following equation:

$$\frac{1}{3}u(\$100) + \frac{2}{3}u(\$200) = \frac{1}{3}u(\$20) + \frac{2}{3}u(\$320) \quad (23)$$

Simple algebraic transformations show that

$$\{u(\$100) - u(\$20)\}/\{u(\$320) - u(\$200)\} = 2 \quad (24)$$

That is, the utility difference between the two amounts of money you could obtain were O_1 to occur is twice as big as the utility difference between the two amounts you could obtain were O_2 to occur. (equation numbers added)

Using phraseology similar to that of Broome (1991, p. 147), Risse (2002, pp. 564) goes on to talk about considerations in favour of one prospect or the other, concluding that

... the utility values tell us how much differences in well-being *count proportionately* in the determination of the comparative overall well-being pertaining to those prospects.

But if the difference in well-being between \$100 and \$20 *counts* for twice as much as the difference between \$320 and \$200 in the determination of your overall well-being, it is plausible to infer that these differences measure *genuine differences* in well-being. (emphasis in the original)

It might seem that Risse is simply endorsing Broome's argument, and I believe that this is what he intended. However, Broome has an independent basis for measuring utility differences (the degree of goodness or well-being). Risse does not. Risse makes inferences about utility differences by *assuming* that only von Neumann-Morgenstern utility functions represent his betterness relation (the move from (23) to (24)), whereas Broome starts with an independent quantitative measure of the degree of well-being.

I am not suggesting that Risse has made the fallacious inference concerning the meaningfulness of intrapersonal utility differences that I discussed in Section 4. Risse explicitly takes account of the convex combination operators in von Neumann and Morgenstern's theory, and this is what allows Risse to regard von Neumann-Morgenstern utility theory as being cardinal, provided that it is accepted that the numerical relational structure is what I have called \mathfrak{N}^4 , which Risse implicitly does. But the choice of numerical relational structure is exactly what is at issue. Unless some compelling justification can be offered for the choice of \mathfrak{N}^4 , not simply that the choice is "natural" or convenient, Risse's argument does not establish that individual well-being is cardinal in the sense required by utilitarianism.

7.3. Broome (1997)

Broome (1997) is concerned with whether a preference-based version of utilitarianism, such as Harsanyi's, is coherent. He ultimately argues that it is not, but this is because he believes that Harsanyi's account of how interpersonal comparisons of goodness are made contains non-preferencist features. Broome now distinguishes between the betterness relation that describes what is good for an individual (and that should form the basis for utilitarian comparisons of alternatives) and this individual's preference relation. For intrapersonal comparisons, Broome identifies the betterness relation with the preferences of an individual in certain ideal circumstances (well informed, calm, etc.) who makes judgments on behalf of herself in her actual nonideal circumstances. For my purposes, we can then simply speak of an individual preference relation.

In his article, Broome adopts the lottery model of uncertainty. Using a variant of his earlier example, but with only three possible sure outcomes, Broome argues for the use of only expectational representations of the individual preference relation. Compared to his analysis in Broome (1991), the link to measurement theory is made more explicit. In Broome's example, (i) sure outcome A is preferred to sure outcome B which in turn is preferred to sure outcome C and (ii) the lottery in which A is received with probability $1/3$ and C is received with probability $2/3$ is indifferent to B for certain.

In singling out the von Neumann–Morgenstern utility functions from the set of all utility functions representing the individual preference relation, Broome (1997, pp. 15–16 in manuscript) says that:

The use of probabilities provides a natural analogue of a pair of scales for measuring the strength—analogue to weight—of preferences. In the example, two chances of the loss from B to C balance the scales against one chance of the gain from B to A , so we naturally take the preference for the gain to be twice as strong as the preference against the loss. The rival concepts are less natural. Compare our concept of physical weight. Any increasing transform of weight could supply a rival concept of weight, but it would be less natural than our present concept. We use our concept because it has the natural and convenient feature that two objects each weighing one pound balance in a scale against one object weighing two pounds.

The use of the weight analogy in this quotation is not particularly apt as the concatenation operator used to measure weight is formally quite distinct from the convex combination operators used in expected utility theory, but this is a minor quibble. What is more important about this quotation is that it appears that, unlike in Broome (1991), this argument does not utilize an independent quantitative measure of degree of preference. Rather, strength of preference is inferred from the expectational representations. Expressed in the language of measurement theory, Broome clearly recognizes that some justification for the choice of numerical relational structure \mathfrak{N}^4 in expected utility theory is required. His justification echoes the one offered by von Neumann and Morgenstern (1944)—this choice is “natural and convenient.” But, as I have argued above, this is an unsatisfactory basis on which to determine the choice of the numerical relational structure in a normative theory.

8. Concluding Remarks

The criticism first raised by Sen (1976) and later formalized by Weymark (1991) that the utility functions that Harsanyi has used in his Aggregation and Impartial Observer Theorems are ordinal and therefore cannot serve a basis for utilitarianism was based on an incomplete description of von Neumann–Morgenstern expected utility theory. Missing from their accounts is the fundamental role that the operators for combining events plays in the von Neumann–Morgenstern representation theorem.

What I have attempted to show here is that even if full account is taken of these operators, von Neumann–Morgenstern utility functions are not cardinal in a normatively compelling way. The problem is that the cardinality of a von Neumann–Morgenstern utility function is dependent on the choice of \mathfrak{N}^4 as the numerical relational structure used to measure utility. The rationales that have been offered for this choice suffice when the objective is to explain individual behaviour, but they do not suffice when the objective is to use utility functions in a normative theory such as utilitarianism. I therefore conclude that Sen’s criticism that von Neumann–Morgenstern utility theory is an inadequate basis for Harsanyi to build a case for utilitarianism is also valid when the more complete description of the empirical relational structure employed by von Neumann and Morgenstern (1944) is taken into account. The interpersonal and intrapersonal utility comparisons that are required for

utilitarianism to be a meaningful doctrine must be found elsewhere.

Given that the main elements of the representational theory of measurement were not systematized until the late 1950s, it is perhaps not surprising that Harsanyi, writing in the early 1950s, appears not to have appreciated the significance for his theory of von Neumann and Morgenstern's remarks that they had "discovered" a natural operation that permitted them to quantitatively measure utility. But he was not alone in this respect. To the best of my knowledge, in his published writings, Harsanyi never explicitly made reference to the ways in which utility differences are quantified in measurement theory. However, in the letter cited in footnote 18 above, Harsanyi draws an analogy between the measurement of utility differences and the measurement of differences in weight. He goes on to describe an operational procedure for measuring weight differences. It is a pity that he never presented these ideas in print, as they provide valuable insight into Harsanyi's views about the issues that I have considered in this article.

Acknowledgements

This article is based on my presentation to the Conference on the History of Social Choice from Condorcet to Arrow and Beyond held at the Université de Caen, October 1–2, 2002. I have benefitted from conversations about measurement theory with Patrick Suppes and Thierry Marchant.

References

- Alchian, A. A., 1953. The meaning of utility measurement. *American Economic Review* 43, 26–50.
- Arrow, K. J., 1951. *Social Choice and Individual Values*. John Wiley & Sons, New York.
- Baumol, W. J., 1951. The Neumann-Morgenstern utility index—an ordinalist view. *Journal of Political Economy* 59, 61–66.
- Baumol, W. J., 1958. The cardinal utility which is ordinal. *Economic Journal* 68, 665–672.
- Berka, K., 1983. *Measurement: Its Concepts, Theories and Problems*. D. Reidel, Dordrecht.
- Blackorby, C., Donaldson, D., Weymark, J. A., 1980. On John Harsanyi's defences of utilitarianism. Discussion Paper No. 8013, Center for Operations

- Research and Econometrics, Université Catholique de Louvain.
- Blackorby, C., Donaldson, D., Weymark, J. A., 1999. Harsanyi's social aggregation theorem for state-contingent alternatives. *Journal of Mathematical Economics* 32, 365–387.
- Bossert, W., Weymark, J. A., 2004. Utility in social choice. In: Barberà, S., Hammond, P. J., Seidl, C. (Eds.), *Handbook of Utility Theory. Volume 2: Extensions*. Kluwer Academic Publishers, Boston, pp. 1099–1177.
- Broome, J., 1991. *Weighing Goods: Equality, Uncertainty and Time*. Basil Blackwell, Oxford.
- Broome, J., 1993. Goodness is reducible to betterness: The evil of death is the value of life. In: Koslowski, P., Shionoya, Y. (Eds.), *The Good and the Economical: Ethical Choices in Economics and Management*. Springer-Verlag, Berlin, pp. 70–84.
- Broome, J., 1997. Can there be a preference-based utilitarianism? In: Salles, M., Weymark, J. A. (Eds.), *Justice, Political Liberalism, and Utilitarianism: Themes From Harsanyi and Rawls*. Cambridge University Press, Cambridge, forthcoming.
- Campbell, N. R., 1920. *Physics: The Elements*. Cambridge University Press, Cambridge.
- d'Aspremont, C., Hammond, P. J., 2001. An interview with John C. Harsanyi. *Social Choice and Welfare* 18, 389–401.
- Domotor, Z., 1979. Ordered sum and tensor product of linear utility structures. *Theory and Decision* 11, 375–399.
- Edgeworth, F. Y., 1881. *Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences*. C. Kegan Paul & Co., London.
- Ellingsen, T., 1994. Cardinal utility: A history of hedonimetry. In: Allais, M., Hagen, O. (Eds.), *Cardinalism: A Fundamental Approach*. Kluwer Academic Publishers, Dordrecht, pp. 105–165.
- Ellis, B., 1966. *Basic Concepts of Measurement*. Cambridge University Press, Cambridge.
- Ellsberg, D., 1954. Classic and current notions of “measurable utility”. *Economic Journal* 64, 528–556.
- Fishburn, P. C., 1989. Retrospective on the utility theory of von Neumann and Morgenstern. *Journal of Risk and Uncertainty* 2, 127–158.
- Fleming, M., 1952. A cardinal concept of welfare. *Quarterly Journal of Economics* 66, 366–384.
- Friedman, M., Savage, L. J., 1948. The utility analysis of choices involving risk. *Journal of Political Economy* 56, 279–304.

- Friedman, M., Savage, L. J., 1952. The expected-utility hypothesis and the measurability of utility. *Journal of Political Economy* 60, 463–474.
- Harsanyi, J. C., 1953. Cardinal utility in welfare economics and in the theory of risk-taking. *Journal of Political Economy* 61, 434–435.
- Harsanyi, J. C., 1955. Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *Journal of Political Economy* 63, 309–321.
- Harsanyi, J. C., 1975. Nonlinear social welfare functions: Do welfare economists have a special exemption from Bayesian rationality? *Theory and Decision* 6, 311–332.
- Harsanyi, J. C., 1977a. Non-linear social welfare functions: A rejoinder to Professor Sen. In: Butts, R. E., Hintikka, J. (Eds.), *Foundational Problems in the Special Sciences*. D. Reidel, Dordrecht, pp. 293–296.
- Harsanyi, J. C., 1977b. *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*. Cambridge University Press, Cambridge.
- Harsanyi, J. C., 1979. Bayesian decision theory, rule utilitarianism, and Arrow’s theorem. *Theory and Decision* 11, 289–317.
- Harsanyi, J. C., 1987. Von Neumann–Morgenstern utilities, risk taking, and welfare. In: Feiwel, G. R. (Ed.), *Arrow and the Ascent of Modern Theory*. Macmillan, London and Basingstoke, pp. 545–558.
- Hernstein, I. N., Milnor, J., 1953. An axiomatic approach to measurable utility. *Econometrica* 21, 291–297.
- Krantz, D. H., Luce, R. D., Suppes, P., Tversky, A., 1971. *Foundations of Measurement. Volume I: Additive and Polynomial Representations*. Academic Press, New York.
- Luce, R. D., Raiffa, H., 1957. *Games and Decisions: Introduction and Critical Survey*. John Wiley & Sons, New York.
- Mandler, M., 1999. *Dilemmas in Economic Theory: Persisting Foundational Problems of Microeconomics*. Oxford University Press, New York.
- Marschak, J., 1950. Rational behavior, uncertain prospects, and measurable utility. *Econometrica* 18, 111–141.
- Mongin, P., d’Aspremont, C., 1998. Utility theory and ethics. In: Barberà, S., Hammond, P. J., Seidl, C. (Eds.), *Handbook of Utility Theory. Volume 1: Principles*. Kluwer Academic Publishers, Boston, pp. 371–481.
- Pareto, V., 1906. *Manuale di Economia Politica, con una Introduzione alla Scienza Sociale*. Società Editrice Libreria, Milan.
- Pfanzagl, J., 1959a. *Die Axiomatischen Grundlagen einer Allgemeinen Theorie des Messens*. Physica-Verlag, Würzburg.
- Pfanzagl, J., 1959b. A general of theory of measurement—applications to

- utility. *Naval Research Logistics Quarterly* 6, 283–294.
- Risse, M., 2002. Harsanyi’s ‘utilitarian theorem’ and utilitarianism. *Noûs* 36, 550–577.
- Roberts, F. S., 1979. *Measurement Theory with Applications to Decision-making, Utility, and the Social Sciences*. Addison-Wesley, Reading, MA.
- Roemer, J. E., 1996a. Harsanyi’s impartial observer is *not* a utilitarian. In: Salles, M., Weymark, J. A. (Eds.), *Justice, Political Liberalism, and Utilitarianism: Themes From Harsanyi and Rawls*. Cambridge University Press, Cambridge, forthcoming.
- Roemer, J. E., 1996b. *Theories of Distributive Justice*. Harvard University Press, Cambridge, MA.
- Savage, L. J., 1954. *The Foundations of Statistics*. John Wiley & Sons, New York.
- Scott, D., Suppes, P., 1958. Foundational aspects of theories of measurement. *Journal of Symbolic Logic* 23, 113–128.
- Sen, A., 1970. *Collective Choice and Social Welfare*. Holden-Day, San Francisco.
- Sen, A., 1976. Welfare inequalities and Rawlsian axiomatics. *Theory and Decision* 7, 243–262.
- Sen, A., 1977. Non-linear social welfare functions: A reply to Professor Harsanyi. In: Butts, R. E., Hintikka, J. (Eds.), *Foundational Problems in the Special Sciences*. D. Reidel, Dordrecht, pp. 297–302.
- Sen, A., 1986. Social choice theory. In: Arrow, K. J., Intriligator, M. D. (Eds.), *Handbook of Mathematical Economics*. Vol. III. North-Holland, Amsterdam, pp. 1073–1181.
- Stevens, S. S., 1946. On the theory of measurement. *Science* 103, 677–680.
- Suppes, P., Zinnes, J. L., 1963. Basic measurement theory. In: Luce, R. D., Bush, R. B., Galanter, E. (Eds.), *Handbook of Mathematical Psychology*. Vol. I. John Wiley & Sons, New York, pp. 1–76.
- Vickrey, W., 1945. Measuring marginal utility by reactions to risk. *Econometrica* 13, 319–333.
- von Neumann, J., Morgenstern, O., 1944. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton.
- Weymark, J. A., 1991. A reconsideration of the Harsanyi–Sen debate on utilitarianism. In: Elster, J., Roemer, J. E. (Eds.), *Interpersonal Comparisons of Well-Being*. Cambridge University Press, Cambridge, pp. 255–320.